UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/695,125	10/28/2003	Manoj Singhal	15153US01	6118

23446 7590 07/01/2010

MCANDREWS HELD & MALLOY, LTD
500 WEST MADISON STREET
SUITE 3400
CHICAGO, IL 60661

EXAMINER	
GODBOLD, DOUGLAS	

ART UNIT	PAPER NUMBER
2626	

MAIL DATE	DELIVERY MODE
07/01/2010	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No. 10/695,125	Applicant(s) SINGHAL, MANOJ	
	Examiner DOUGLAS C. GODBOLD	Art Unit 2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *29 April 2010*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-5,9-15 and 28* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-5,9-15 and 28* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

1. This Office Action is in response to correspondence filed April 29, 2010 in reference to application 10/695,125. Claims 1-5, 9-15 and 28 are pending and have been examined.

Response to Arguments

2. Applicant's arguments filed April 29, 2010 have been fully considered but they are not persuasive.

3. Regarding applicant's arguments, see Remarks page 2 and 3, that the combination of Saunders and Tzanetakis fail to teach "analyzing selected audio frequency components," the examiner respectfully disagrees. Saunders, as discussed in the previous rejection, discusses analyzing time domain audio components. Tzanektakis, however, teaches taking audio frequency components, and either converting them to time frequency components (see introduction) or using features from the frequency domain to do the segmentation (see sections 2 and 3). Thus Tzanektakis teaches analyzing audio frequency components.

4. Regarding applicant's arguments that Saunders teaches away from the combination, the examiner respectfully disagrees. Applicant contends that Saunders teaches away from using frequency features, because Saunders discusses that ZCR, a time domain feature, is a good music/speech discriminator. The examiner notes,

however, that Saunders does not state that this is the only speech discriminator. Moreover, Saunders is silent on what to do in the case were the data is store in the frequency domain already (such as MPEG). Tzanektakis provides a solutions for when audio signals are in the frequency domain… analyze frequency domain features (see introduction). Therefore Saunders and Tzanektakis are combinable.

Claim Rejections - 35 USC § 103

5. The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

6. Claims 1, 3-5, 10, 11-13, 14, and 28 are rejected under 35 U.S.C. 103(a) as being unpatentable over Saunders (Real-Time Discrimination of Broadcast Speech/Music) in view of Tzanetakis et al (Sound analysis Using MPEG Compressed Audio) in view of Yamato (US PAP 2004/0193406).

7. Consider claim 1, Saunders teaches a method for classifying an audio signal (we describe a technique which is successful at discriminating speech from music; page 993, column 1, line 1), the method comprising:

receiving an audio signal to be classified (this is a technique for discriminating speech from music from an FM broadcast; page 993, column 1, line 2);

analyzing selected audio signal components (The first step is to measure the ZCR of the signal over a 2.4 second segment of the data; page 994, column 2, line 43);

recording a result of analysis of the selected audio signal components (would be inherent in order to compare it);

comparing the recorded result of analysis to a threshold value (If this statistic exceeds a specific threshold, the distribution outside these bounds is significantly skewed and the waveform is likely speech; page 994, column 2, line 43); and

classifying the audio signal based upon comparison of the recorded result of analysis and the threshold value (If this statistic exceeds a specific threshold, the distribution outside these bounds is significantly skewed and the waveform is likely speech; page 994, column 2, line 43).

Saunders does not specifically teach that the audio signal components are audio frequency components.

In the same field of audio analysis, Tzanetakis teaches the audio signal components are audio frequency components. (Tzanetakis analyzes MPEG audio files, which stores subband information that were converted via a filter bank; overview of MPEG 2. In order to use traditional analysis such as zero-crossing, MPEG data must be decoded; introduction 1, paragraph 2. Tzanetakis offers the solution of using frequency features instead; see sections 2 and 3).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to combine the use of frequency vectors of Tzanetakis with system of Saunders in order to avoid having to convert MPEG files to time domain in order to apply traditional analysis, Tzanetakis introduction 1, paragraph 2.

Saunders and Tzanetakis do not specifically teach selected audio frequency components having a frequency less than a predetermined frequency in the speech range.

In the same field of signal classification, Yamato teaches selecting audio frequency components having a frequency less than a predetermined frequency (Figure 4, signal is preprocessed, removing components above 4kHz, paragraph 0062).

Therefore it would have been obvious to combine use the filtering in Yamato with the system of Saunders and Tzanetakis to remove extraneous noise that may interfere with signal classification (Yamato, paragraph 0062)

8. Consider claim 3, Saunders and Tzanetakis teach the method according to claim 1, wherein analyzing the selected audio frequency components comprises transforming the selected audio frequency components to time domain components (Tzanetakis analyzes MPEG audio files, which stores subband information that were converted via a filter bank; overview of MPEG 2. decoding MPEG inherently comprises transforming subband information back to time domain signals) and counting zero point transitions of the selected audio signal components (The first step is to measure the ZCR of the signal over a 2.4 second segment of the data; page 994, column 2, line 43. Measuring the Zero Crossing Rate would entail counting the number of zero crossings).

9. Consider claim 4, Saunders and Tzanetakis teach the method according to claim 1, wherein recording a result of analysis of the selected audio frequency components

comprises transforming the selected audio frequency components to time domain components (Tzanetakis analyzes MPEG audio files, which stores subband information that were converted via a filter bank; overview of MPEG 2. decoding MPEG inherently comprises transforming subband information back to time domain signals) and recording a count value of a number of zero point transitions of the selected audio signal components (The first step is to measure the ZCR of the signal over a 2.4 second segment of the data; page 994, column 2, line 43. Measuring the Zero Crossing Rate would entail counting the number of zero crossings. This number would inherently have to be stored somewhere in order to process it or manipulate it).

Consider claim 5, Saunders, Tzanetakis and Yamato the method according to claim 1, further comprising selecting audio frequency components prior to analyzing selected audio frequency components, wherein said selecting audio frequency components comprises passing the audio signal through a low pass filter for filtering out audio frequency components having a frequency greater than a predetermined frequency thereby reducing an amount of audio information to be analyzed. (Figure 4, signal is preprocessed, removing components above 4kHz, paragraph 0062. This is a lowpass filter).

10. Consider claim 10, Saunders teaches the method according to claim 1, wherein classifying the audio signal occurs at a receiving end of an audio transmission system

(this is a technique for discriminating speech from music from an FM broadcast; page 993, column 1, line 2).

11. Consider claim 11, Saunders teaches the method according to claim 1, wherein the audio signal is one of an analog signal and a digital signal (A sample rate of 16Khz was chosen for this discrimination technique; page 995, column 1 line 1. If something is sampled it is well understood that it is being converted to a digital signal. this is a technique for discriminating speech from music from an FM broadcast; page 993, column 1, line 2. This further tells us that the signal started out as an analog signal as at the time of the publication of Saunders all FM broadcasts were analog.).

12. Consider claim 12, Saunders, Tzanetakis, and Yamato teach the method according to claim 1, but does not specifically teach wherein the threshold value used in the comparison is pre-determined and pre-set by a user.

However Saunders does teach Data was collected manually by listening, collecting and storing features, and labeling the segment. A variety of content was processed, including talk, commercials, and many types of music. Once the classifier was trained, the parameters were stored and fed into the real-time feature extraction/classifier routine; page 995, column 1, line 33.

With data being collected manually, it must be entered manually, and although is not specifically the threshold, one of ordinary skill in the art that the training of the

classifier by manually collecting data is changing the threshold. Therefore in fact, the user is in a way changing the threshold value is preset and determined by the user.

13. Consider claim 13, Saunders teaches the method according to claim 1, wherein the threshold value used in the comparison determined through trial and error of a plurality of iterations in a comparing device (Data was collected manually by listening, collecting and storing features, and labeling the segment. A variety of content was processed, including talk, commercials, and many types of music. Once the classifier was trained, the parameters were stored and fed into the real-time feature extraction/classifier routine; page 995, column 1, line 33).

14. Consider claim 14, Saunders teaches the method according to claim 1, wherein analyzing selected audio frequency components comprises counting zero point transitions of the audio signal for a predetermined period of time (The first step is to measure the ZCR of the signal over a 2.4 second segment of the data; page 994, column 2, line 43. Measuring the Zero Crossing Rate would entail counting the number of zero crossings).

15. Consider claim 28, Yamato teaches the method according to claim 1, wherein the predetermined frequency is approximately 4K Hz (Yamato, paragraph 0062, frequencies above 4KHz removed).

16. Claim 2 is rejected under 35 U.S.C. 103(a) as being unpatentable over Saunders in view of Tzanetakis in view of Yamato as applied to claim 1 above, and further in view of Carey (A Comparison of Features for Speech, Music Discrimination).

17. Consider claim 2, Saunders in view of Tzanetakis and Yamato teaches the method according to claim 1, but does not specifically teach wherein classifying the audio signal based upon comparison of the recorded result of analysis and the threshold value further comprises:

if the recorded result of analysis is greater than the threshold value, then the audio signal is determined to be music; and

if the recorded result of analysis is less than the threshold value, then the audio signal is determined to be speech.

In the same field of speech/music discrimination, Carey teaches if the recorded result of analysis is greater than the threshold value, then the audio signal is determined to be music (table 1 shows that the mean value of number of zero crossing (u) for music 0.18 is greater than that of speech 0.17); and

if the recorded result of analysis is less than the threshold value, then the audio signal is determined to be speech (table 1 shows the mean value of zero crossing for speech 0.17 was less that music 0.18).

Although Saunders in view of Tzanetakis and Yamato uses a slightly different zero crossing analysis method than does Carey, it would have been obvious to one of

ordinary skill in the art at the time of the invention to use the parameters of Carey as this method would be computationally inexpensive (Carey page 151, column 2, section 4.4).

18. Claims 9, 15 are rejected under 35 U.S.C. 103(a) as being unpatentable over Saunders and in view of Tzanetakis and Yamato as applied to claim 1 above and further in view of Benyassine.

19. Consider claim 9, Saunders and Tzanetakis and Yamato teach the method according to claim 1, but does not teach specifically wherein classifying the audio signal occurs at a transmitting end of an audio transmission system.

However in the same field of music and speech discrimination Benyassine teaches classifying the audio signals at a transmitting end of an audio transmission system (Figure 1, encoder 112, part of transmission side, may contain a music classifier with voice activity detector; column 3, line 62.)

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to classify the music or voice at the transmitting side of the system as taught by Benyassine in order to determine properties of the signal in order to best encode the signal for transmission (Benyassine; column 1 line 62 - column 2 line 13).

20. Consider claim 15, Saunders Tzanetakis and Yamato teach the method according to claim 1, but does not specifically teach further comprising:

converting the audio signal from an analog signal to a digital signal;

encoding the audio signal;

packetizing the audio signal;

transmitting the audio signal;

decoding the audio signal; and

processing the audio signal, wherein processing at least comprises one of storing the audio signal and playing the audio signal.

However in the same field of music and speech discrimination Benyassine teaches converting the audio signal from an analog signal to a digital signal (figure 1, A/D converter 108);

encoding the audio signal (figure 1, encoder 112);

packetizing the audio signal (communication devices 102 and 106 may be cellular telephones radios, or VoIP systems; column 3 line 6-11. Cell phones and VoIP systems both used packetized data);

transmitting the audio signal (figure 1, signals are transmitted over communication medium 104);

decoding the audio signal (using decoder 114, figure 1); and

processing the audio signal, wherein processing at least comprises one of storing the audio signal and playing the audio signal (output of system is synthesized speech signal 120, figure 1).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the transmission scheme of Benyassine with the audio classification method of Saunders, Tzanetakis, and Yamato in order to provide an

efficient way to effectively transmit audio signals (Benyassine; column 1 line 62 - column 2 line 13).

Conclusion

21. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to DOUGLAS C. GODBOLD whose telephone number is (571)270-1451. The examiner can normally be reached on Monday-Thursday 7:00am-4:30pm Friday 7:00am-3:30pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (571) 272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

DCG

/Richemond Dorvil/
Supervisory Patent Examiner, Art Unit 2626